Copy 2 of 2

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	4/24/2002	Final Technical Report

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
A Cooperative RSTA Reseach Program in Software and Intelligent Systems	DAAH04-96-1-0429
6. AUTHOR(S) Dr. R. Bajcsy	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Pennsylvania/Computer & Info Sci GRASP 3401 Walnut ST, Rm 303 C Phila., PA 19104	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSORING / MONITORING AGENCY REPORT NUMBER 35798.1-CI

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT	12 b. DISTRIBUTION CODE
Approved for public release; distribution unlimited.	

13. ABSTRACT (Maximum 200 words)

See attached.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

Enclosure 1

20021017 057

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. NOT TO BE USED FOR INTERIM PROGRESS REPORTS; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

U.S. Army Research Office
ATTN: AMSRL-RO-BI (TR)
P.O. Box 12211
Research Triangle Park, NC 27709-2211

- [ ] Reprint (Orig + 2 copies)
- [ ] Manuscript (1 copy)
- [ ] Technical Report (Orig + 2 copies)
- [x] Final Progress Report (Orig + 2 copies)
- [ ] Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER: DAAH04-96-1-0429

REPORT TITLE:
A COOPERATIVE RSTA RES. PGM. IN SOFTWARE & INTELLIGENT SYSTEMS."
is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

Max Mintz

MAX MINTZ
UPENN CIS

A Cooperative RSTA Research Program in Software and Intelligent Systems

Army Research Office Grant DAAH04-96-1-0429

Final Technical Report
Reporting Period: 01 SEP 96 through 30 NOV 98

Professor Ruzena Bajcsy, PI
GRASP Laboratory
3401 Walnut Street, Room 303C
University of Pennsylvania
Philadelphia, PA 19104-6228

Manuscripts:

[1] R. Bajcsy and E. W. Large. "When and Where will AI Meet Robotics: Issues in Representation." *AI Magazine*. Fall 1999.

[2] R. A. Hicks and R. Bajcsy. "Catadioptric Sensors that Approximate Wide-angle Perspective Projections." GRASP Lab Technical Report 1999.

Scientific Personnel Supported:

R. A. Hicks, Post-doctoral Fellow, 01 SEP 96 through 31 AUG 97.

E. W. Large, Post-doctoral Fellow, 01 SEP 97 through 31 AUG 98.

Scientific Progress:

The Integration of Perception-Action Systems with Systems that Reason Abstractly:

Because perception action systems are necessarily constrained by the physics of time and space, robotocists often assume they are best described using differential equations, a language that is specialized for describing the evolution of variables that represent physical quantities. However, when it comes to decision making, where the representation envolved refer to goals, strategies and preferences, AI offers a diverse range of formalisms to the modeler. However, the relationship between these two levels of representation — signal and symbol — are not well understood. If we are to achieve success in modelling intelligent physical agents, robotics and AI must reach a new consensus on how to integrate perception-action systems with systems designed for abstract reasoning. Research towards this end is described in reference [1] by Bajcsy and Large.

Novel Vision Sensors that Approximate Wide-Angle Perspective Projections:

Reference [2] by Hicks and Bajcsy, present two families of reflective surfaces that are cabable of providing a wide field of view, and yet still approximate a perspective projection to a high degree. These surfaces are derived by considering a plane perpendicular to the axis of the surface of revolution and finding the equations governing the distortion of the image of the plane in this surface. We then view this relation as a differential equation and prescribe the distortion term to be linear. By choosing appropriate initial conditions for the differential equation and solving it numerically, we derive the surface shape and obtain a precise estimate as to what degree the the resulting sensor

can approximate a perspective projection. Thus, these surfaces act as computational sensors, allowing for a wide-angle perspective view of a scene without processing the image in software. The applications of such a sensor should be numerous, including surveillance, robotics, and traditional photography.

When and Where Will AI Meet Robotics?

Issues in Representation

Ruzena Bajscy and Edward W. Large

■ Because perception-action systems are necessarily constrained by the physics of time and space, robotocists often assume they are best described using differential equations, a language that is specialized for describing the evolution of variables that represent physical quantities. However, when it comes to decision making, where the representations involved refer to goals, strategies, and preferences, AI offers a diverse range of formalisms to the modeler. However, the relationship between these two levels of representation—signal and symbol—are not well understood. If we are to achieve success in modeling intelligent physical agents, robotics and AI must reach a new consensus on how to integrate perception-action systems with systems designed for abstract reasoning.

In the early days of AI, robotics was an integral part of our research effort. All our major AI laboratories had research programs in robotics in the late 1960s and early 1970s. However by the 1980s, robotics had taken its own course separate from the core activities of AI. One might argue that such a split was inevitable, a natural result of specialization in a rapidly growing and maturing field such as ours, but in our pursuit of rational models of the mind, do we dare leave the body behind?

What is responsible for the divergence between these two fields that once were so intimately intertwined? Can AI and robotics ever be reunited, and if so, what would a new partnership look like? At the core, we believe, is the ubiquitous issue of representation. There is an enormous difference between dealing with physical systems that operate in our everyday environment and software systems that reside in various abstract worlds. This gap has led to divergence in many areas, including the following:

The problems: Robotics problems entail systems and agents interacting with the physical world, but AI deals mostly with abstract problems that lend themselves to symbolic representations.

The environment: Roboticists seek to design systems that function in physical environments that are always changing and intrinsically unpredictable. Software agents generally operate in human-designed worlds where one can have some measure of control over change or at least an a priori knowledge of the possibilities.

The tools: AI more commonly uses discrete mathematics, but robotics and machine perception make use of continuous mathematics. These tools also differentiate the typical educational background that characterizes the two areas: AI has more computer scientists, but robotics has more electrical and mechanical engineers.

The evaluation criteria: AI researchers seem to value novelty more, solving "hard" problems, showing existence of solutions, and so forth. Robotics, however, follows traditional engineering evaluation criteria: efficiency, reliability, accuracy of performance, and economy of the solution.

We admit that these divisions might be overexaggerated because in both fields, one can find counterexamples to the previous statements. Nevertheless, each of these points speaks to the differences one encounters when dealing with corporeal agents in the physical world versus software agents in cyberspace. Robotics concentrates most of its resources on modeling perception and action. Often, differential equations are used to embody relatively simple strategies for controlling hardware effectors based on sensory information. AI, however, emphasizes planning and abstract

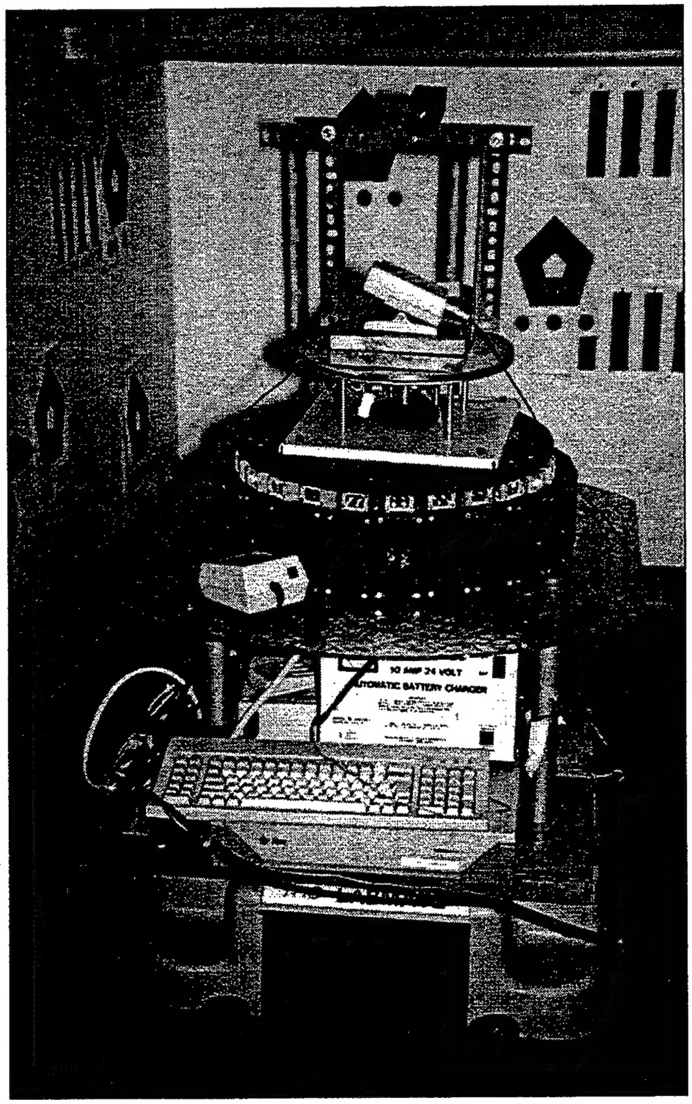

Figure 1. The Autonomous Agent.

The TRC platform serves as a mobile base. A stereo camera pair mounted on the front of the rig is used for obstacle detection. A third camera mounted on a pan platform is used for target detection and tracking.

reasoning. For example, logical, grammatical, or other discrete formalisms are used to model the complex operations involved in winning a chess match or parsing a sentence.

What seems clear is that as robotic agents are called on to perform increasingly complex tasks, they will be required not only to react flexibly in dynamically changing environments but also to make decisions, reason abstractly, and change perceptual or behavioral strategies. Conversely, as intelligent software agents are required to operate more and more on human terms, responding to sensory information and interacting physically with humans, they will be required to integrate more sophisticated perception-action capabilities with their abstract reasoning abilities. In addition, although each of these areas has been well studied, in robotics and AI respectively, the integration of perception-action systems with reasoning systems is less well understood. Thus, there is a great need to reconsider the relationship between AI and robotics.

The Problem of Representation

We will explore the issue that we believe is key to this relationship, the issue of representation. Representation is critical, especially when one considers how to find a description that is compact, yet expressive enough to enable the modeling of intelligent physical agents. What kind of mathematical tools are available to us? At the signal level, modeling of perception-action (reactive) behaviors can be anchored in differential equations and control theory, both linear and nonlinear. At the symbol level, higher-level control derives its models either from geometry (typically used in robotics) or from logics and rule-based systems such as are favored in the AI planning community. If time needs to be explicitly accounted for, then there are other tools available. At the signal level, time is implicit in the model of the reactive behaviors (for example, using differential equations). At the symbol level, discrete states are generally considered, and these can be modeled using discrete-event systems; temporal logics; and, at an even higher level, fluents. If uncertainty and disturbances must be modeled, then one must bring to bear stochastic models and probability theory; partially observable Markov decision models is one such example. Finally, utility functions and cost-benefit trade-offs come to play in conjunction with game theory, optimization, selection of strategies, and complexity considerations. Examples of such approaches as they have been applied to robotics are presented in Alami et. al. (1998); Clementia, Di Felice, and Hernandez (1997); Cohn (1995); Russell and Subramanian (1995); and Sandewall (1994). Hybrid system approaches such as discussed in Arkin (1998), Brockett (1993), Dickmanns (1997), Nagel and Haag (1998), Nerode and Remmel (1991), and Ramadge and Wonham (1987) combine discrete systems with lower-level control systems. However, the use of all these mathematical tools is predicated on the assumption that label assignment (the segmentation of the sensory or control signals) is performed externally to the system.

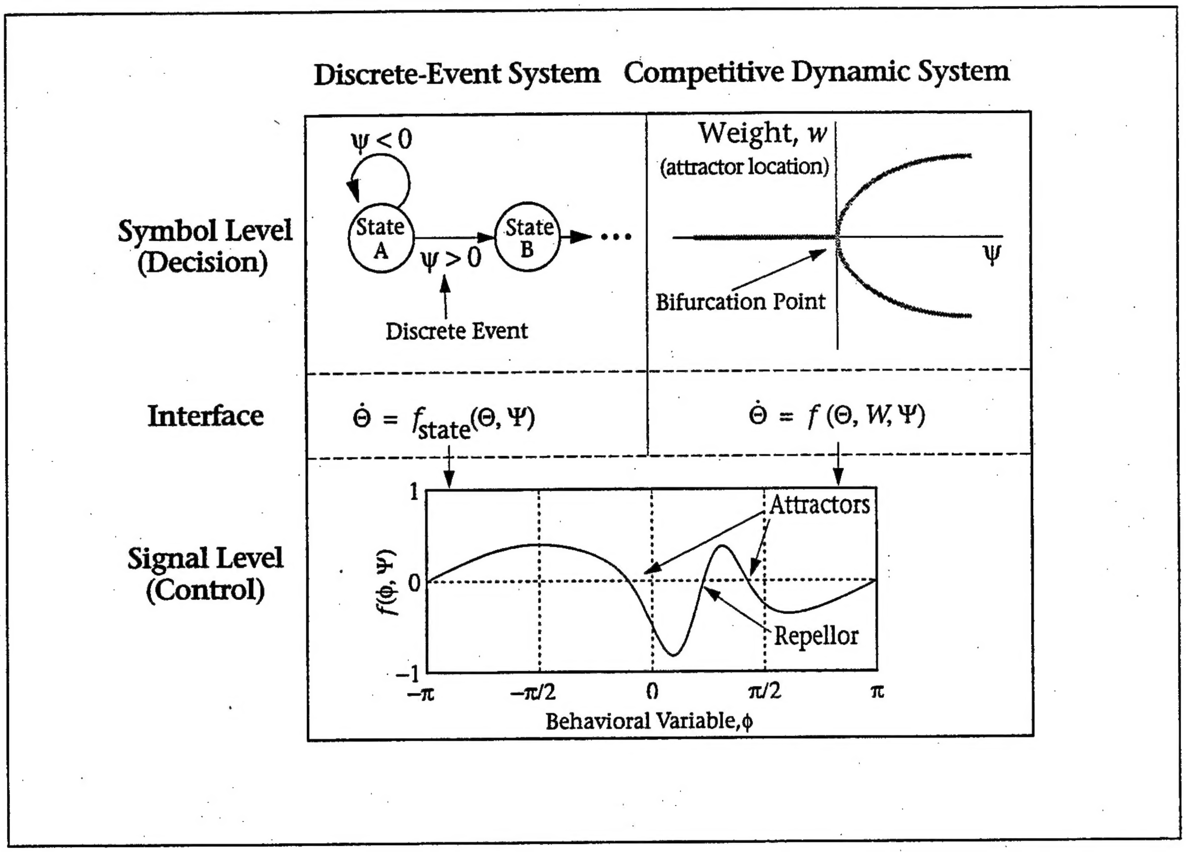

Figure 2. Two Approaches to Decision Making and Behavioral Sequencing.

In the discrete-event system, the discrete states of a finite-state machine correspond to control of perception-action by a distinct control law ($\dot{\Theta} = F_{state}(\Theta, \Psi)$). Transitions between states are governed by guard conditions on perceptual variables (for example, ψ). In the competitive dynamic system, each state variable (for example, w) controls the weighting of a task constraint at the signal level. Behavior is shaped directly by the competitive dynamic system, $\dot{\Theta} = F(\Theta, W, \Psi)$, as the symbol-level system activates and deactivates attractor and repellor contributions to the behavioral dynamics. As perceptual parameters (for example, ψ) change, bifurcations cause qualitative changes in perception-action behavior.

Signal and Symbol

In our view, the main unsolved problem is how to segment the continuous signal into states, strategies, labels, and so forth, and how to arbitrate among states for effective interaction with the environment (for example, Shi and Malik [1998], Tari, Shah, and Pien [1997], and Large, Christensen, and Bajscy [1999]). In the GRASP Laboratory, we have concerned ourselves with the problem of representation in signal-symbol systems over the past several years. One approach to this problem involves the study of *intelligent physical agents,* agents that can operate in the physical world with all its uncertainty yet behave intelligently, making decisions about how best to perform simple and complex tasks in a range of real-world environments. A picture of one such agent is shown in figure 1. It consists of a TRC LABMATE mobile platform equipped with a stereo camera pair used for obstacle detection; a third camera mounted on a turntable for visual tracking; and several computers used for processing sensory signals, generating control signals, and making decisions.

The control system that models perception-action behavior transforms visual input into control signals to enable the physical agent to carry out various navigation tasks. We use a dynamic system approach proposed by Schoner, Dose, and Engels (1996). In this approach, behavior is controlled by differential equations that describe the rate of change of *behavioral variables,* such as heading direction and velocity. At any instant in time, the values of these variables describe the agent's behavior. Over time, the dynamic system generates a series of values, controlling the behavior of the agent. Our dynamic system has the form

Although the dynamic control system provides a great deal of flexibility, it can only model one relatively simple perception-action behavior at a time. Complex tasks, however, typically require the execution of sequences of behavior.

$$D\Theta/dt = F(\Theta,\Psi) \qquad (1)$$

where $\Theta = [\phi\ \upsilon]^T$ is a vector of behavioral variables, heading direction, and velocity and is a vector of variables that represent perceptual information, such as direction to the target and size of obstacles. An example of such a function is shown in figure 2 (bottom panel). Three *fixed points* can be seen in the figure as points where the value of $f(\Theta,\Psi)$ is zero (that is, $d\phi/dt = 0$, so heading direction is fixed). If the slope of the function around a fixed point is positive, the value of the behavioral variable is pushed away from this value by the action of equation 1; such an unstable fixed point is called a *repellor*. If the slope of the function is negative, it is a stable fixed point, called an *attractor*, because the behavioral variable is pulled toward this value by the differential equation.

The behavior of the agent is controlled by the configuration of attractors and repellors: Desired actions (such as moving toward a target) are modeled as attractors, and undesired actions (such as moving toward an obstacle) are modeled as repellors of the perception-action dynamic system. *Task constraints* determine the mapping from perceptual information to behavioral attractors and repellors. If the task is to go to the desk, the action of moving toward the desk is modeled an attractor, but other objects are considered obstacles (modeled as repellors). However, if the task is to rendezvous with another agent, then the action of moving toward the other agent is an attractor, and the desk is treated as an obstacle, and moving toward it is to be avoided. Thus, viewed as a representation of a perception-action behavior, this dynamic system incorporates task knowledge and makes use of perceptual information.

As the values of the perceptual variables change, the attractor-repellor layout changes. For example, if a target moves, the location of the corresponding attractor will move as well, thus providing behavioral flexibility—behavior adapts to accommodate changes in the environment. An even greater measure of flexibility is provided by *bifurcations* in the dynamic system—qualitative changes in the layout of attractors and repellors caused by changes in parameter values. For example, when a new obstacle comes into view, a repellor forms where there was no repellor before.

Although the dynamic control system provides a great deal of flexibility, it can only model one relatively simple perception-action behavior at a time. Complex tasks, however, typically require the execution of sequences of behavior. For example, one agent might need to rendezvous with an agent and then proceed toward a target location. To perform such tasks in a complex environment requires sequencing several simpler perception-action strategies. Building on the previous signal-level representation, we have investigated two approaches to modeling decision making and sequence generation: (1) a *discrete-event–system* approach and (2) a *competitive dynamic system* approach. There are three key differences between these two approaches: (1) the model of how the symbol level interfaces with the signal level, (2) the model of how perception is integrated into the decision-making process, and (3) the model of how decisions are captured at the symbol level.

Before describing each approach in detail, we summarize the differences between the two systems in figure 2. In the *discrete-event– system* approach (Kosecka 1996), the symbol level interfaces with the signal level by realizing distinct behaviors as separate dynamic systems (that is, f_{state}, middle left panel); these are conceived of as elementary perception-action strategies. At any particular time, the behavior of the agent is governed by one of these equations. Decision making and sequencing are modeled at the symbol level using a finite-state machine (FSM) (top left panel). Individual FSM states correspond to elementary signal-level behaviors; when in a particular state, behavior is governed by a corresponding signal-level dynamic system. The arcs linking states are labeled with discrete events (for example, $\alpha > 0$) that summarize perceptual information. The perceptual system generates these events, which correspond to qualitatively different conditions in the environment. The occurrence of a specific event causes the switch from one state to another, modeling the decision to execute a different perception-action behavior. Thus, sequences of behavior are generated by traversing the arcs, which is, in turn, governed by the conditions on the current situation.

The *competitive dynamic system* model is formulated entirely within the qualitative theory of dynamic systems. Both signal-level control and symbol-level decision making are modeled using differential equations (Large, Christensen, and Bajscy 1999). However, the competitive dynamic system interfaces with the signal level differently than the discrete-event system. Rather than defining multiple elementary perception-action behaviors, a single master equation is defined containing all possible task constraints (middle right panel, figure 2). Then, each variable in the symbol-level system controls the weighting of a task constraint at the signal level, such that the symbol-level sys-

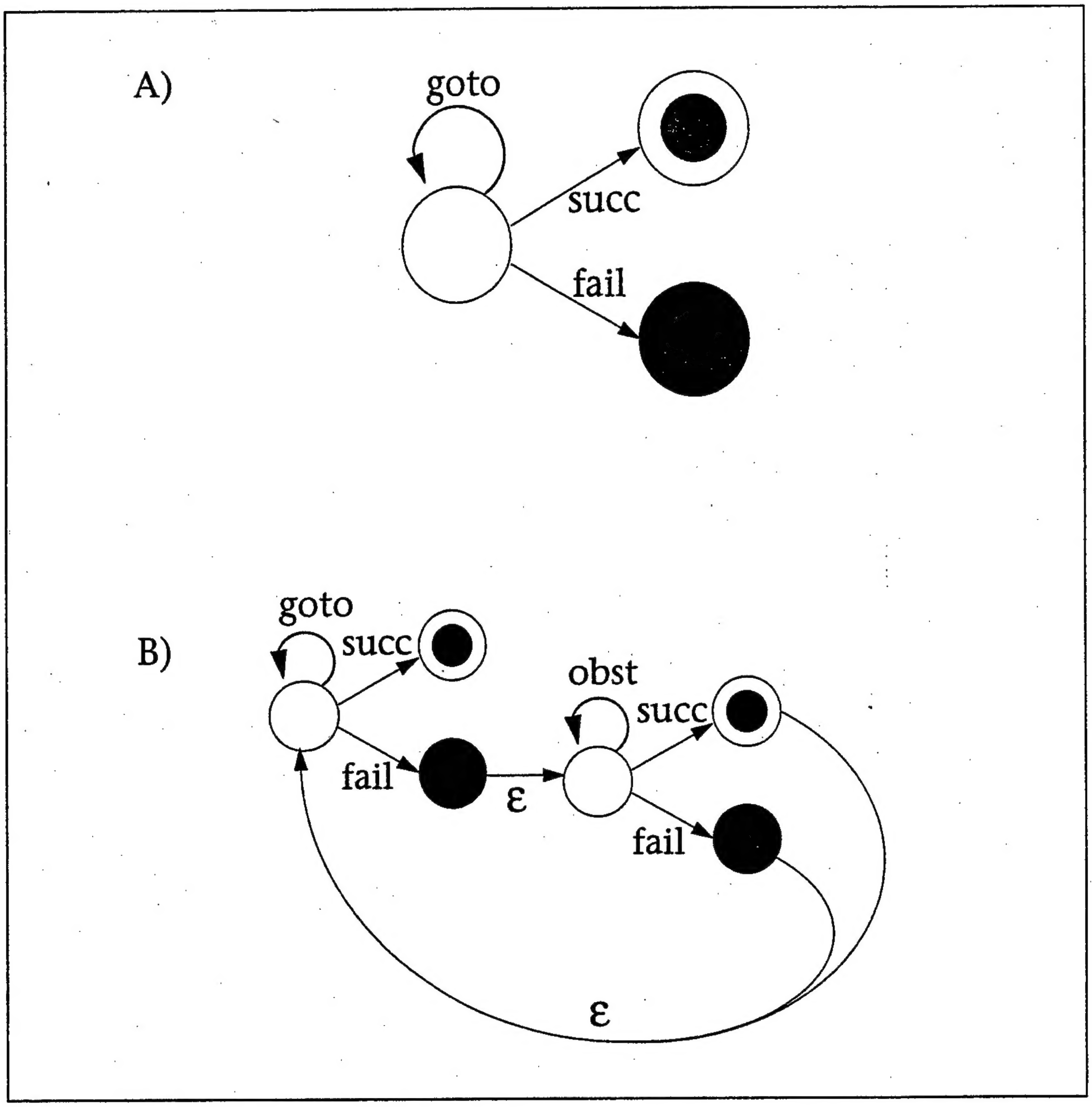

Figure 3. Finite-State Machine (FSM) *Models for Simple and Complex Behaviors.*

A. An FSM for elementary behavior GoTo. The control law (f_{GoTo}) is repeatedly invoked in the next state until successful (arrival at the goal) or unsuccessful (for example, detection of a spurious attractor) termination.

B. Finite-state model for a navigation behavior. Failure of GoTo is followed by the elementary behavior Escape. Once the agent clears the obstruction, GoTo is invoked again. This more complex behavior is able to handle a large variety of navigation situations.

tem can activate and deactivate attractor and repellor contributions to the behavioral dynamics. Different behaviors are modeled as fixed points (stable weight configurations) in the competitive dynamic system. The environment determines the values of the system parameters. As the perceptual information changes, parameters change, causing bifurcations in the symbol-level system (top right panel, figure 2), modeling the decision to cease executing one behavior and execute another instead. This approach differs from the discrete-event approach because the properties of dynamic systems, such as stability, bifurcations, and hysteresis, govern decision making and sequence generation.

Discrete-Event Systems

The discrete-event approach models elementary perception-action strategies as behavioral atoms, and each elementary control law is associated with a state in a simple finite-state machine. We define the composition operators for the FSMs by imposing some additional structure. The set of final states of an elementary behavior is partitioned into a set of successful and unsuccessful final states. By utilizing these primitives, it is possible to build

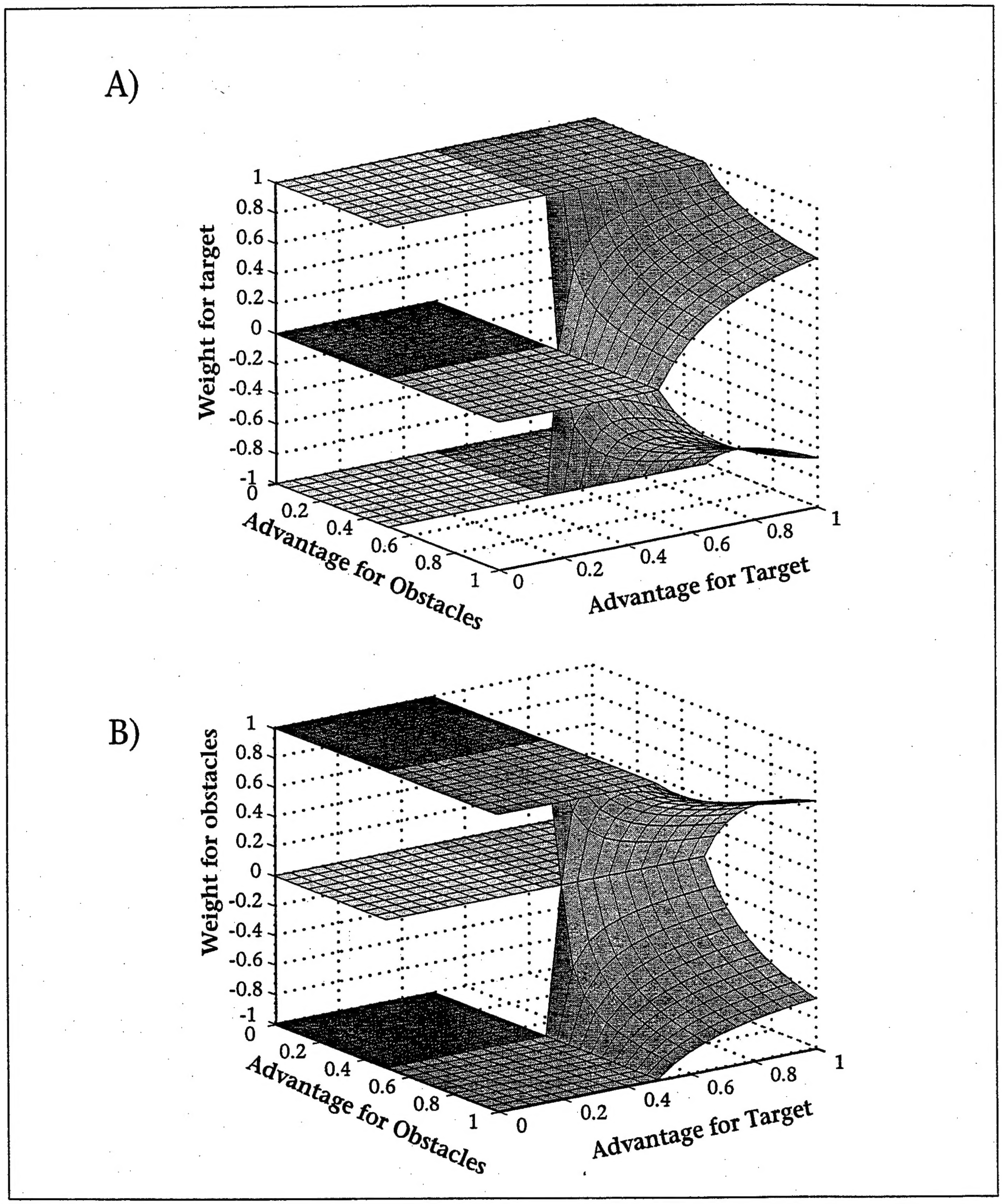

Figure 4. Bifurcations in a Competitive Dynamic System for Decision Making.

Two bifurcation diagrams are shown, one corresponding to each dimension of a two-dimensional system for system navigation. It is assumed that $\gamma_{tar,obs} = \gamma_{obs,tar} = 0.5$. A. The state variable w_{obs} determines the weighting of f_{tar}. B. The state variable w_{obs} determines the weighting of f_{obs}. Four qualitatively different behaviors, corresponding to four fixed points of the competitive dynamics, are shown.

models of more complex tasks, which are described as sequences of behavioral atoms. By using a task-specification language, complex FSMs are synthesized by sequencing simpler automata.

The FSM model of an elementary GoTo strategy is shown in figure 3a. A signal-level perception-action behavior is repeatedly invoked in the next state until the agent reaches the desired target location and makes a transition to the final state. If the strategy fails, the transition to the unsuccessful final state is made. As an example of a composition of elementary behaviors, consider the problem of moving to a target location while avoiding obstacles. In simple environments, one elementary perception-action behavior, GoTo, might do the trick. However, in complex environments with multiple obstacles arrayed in difficult configurations, our agent might get stuck in an area and never reach the target location (Large, Christensen, and Bajscy 1999). We address this problem by adding an Escape behavior that enables the agent to find its way out of enclosures and other spatial traps. The FSM for this behavior is shown in figure 3b. We assume that the fail signal to the GoTo behavior is generated whenever the agent detects an enclosure from which it must escape. When the fail signal is generated, the FSM enters the unsuccessful final state, and a transition is made to the initial state of Escape. When Escape terminates (when the agent has cleared the obstruction), the GoTo behavior resumes. The navigation task terminates successfully when the agent reaches the target location.

Competitive Dynamic Systems

The competitive dynamic system strategy models individual behaviors as the stable fixed points of a decision-making dynamic system. This system interacts with the signal level not by invoking separate elementary behaviors but by directly shaping perception-action strategies. The variables of the competitive dynamic system determine the weighting of the task constraints in the behavioral dynamic system. This interface between the two levels allows the decision-making system to activate and deactivate attractors and repellors in the perception-action system, synthesizing control laws on the fly. Thus, qualitatively different configurations of the weights give rise to distinct perception-action behaviors. In addition, distinct weight configurations, which arise as attractors in the competitive dynamic system, are functionally equivalent to the FSM states of the discrete-event system.

Decisions are made through bifurcations in the competitive dynamic system, and as with any dynamic system, bifurcations are caused by changes in the values of the system parameters. The decision-making system uses two types of parameter: (1) competitive advantage and (2) competitive interaction. These parameters are tied to perceptual information, so that decisions are made on the basis of the environment as sensed by the agent. First, each weight has an associated *competitive advantage* that describes whether the corresponding task constraint (for example, move toward target, avoid obstacles) is appropriate to the agent's current situation. For example, if obstacles are nearby, the Obstacles constraint will have a strong competitive advantage, but if the target is also in view, the Target constraint also has a strong advantage. The activation of both constraints simultaneously corresponds to the elementary GoTo behavior of the discrete-event system earlier. *Competitive interaction* describes the extent to which each constraint is consistent or inconsistent other constraints. For example, if the agent finds itself enclosed in an area with the target just beyond, the competitive interaction between the Obstacles and the Target constraints would increase so that the Target constraint would be deactivated temporarily, allowing the agent to escape from the enclosure. Deactivation of the target constraint corresponds to the Escape behavior of the discrete-event system.

We can understand in detail how these parameters interact to determine the behavior of the agent by constructing a bifurcation diagram such as that in figure 4. The bifurcation diagram shows the layout of fixed points for a two-dimensional system (that is, weights for the Target and Obstacles constraints) as a function of the perceptual parameters. In the figure, the competitive advantage parameters are varied from 0 to 1, assuming that the two competitive interaction parameters remain fixed at 0.5. Four qualitatively different regions (and, thus, behaviors) are pictured. Activation of both constraints corresponds to the GoTo behavior described earlier, but activation of the Obstacles constraint only corresponds to the Escape behavior.

Beginning in the front left corner of the parameter space, only the Obstacles constraint contributes to the behavioral dynamics (the Escape behavior). Moving to the right, as Target's advantage increases beyond 0.5, both Target and Obstacles constraints contribute to shape perception-action behavior (the GoTo behavior). As we next decrease the advantage of Obstacles, moving to the back left region, Obstacles is deactivated, but Target is activat-

Intelligent agents must be capable of bringing to bear a rich variety of perception-action strategies but, at the same time, reasoning and solving problems to perform both familiar and unfamiliar tasks in novel environments.

ed. If we once again decrease Target, moving toward the back left region of the parameter space, we notice that Obstacles remains inactive, but Target remains active. This last region of the state space is different from the others. Two behaviors are stable in this region. However, the system can only occupy one of these states at any given time. In this case, the state of the system is determined by its recent history, a phenomenon known as *hysteresis*. Finally, the boundaries of the four regions are determined by the values of the competitive interaction parameters. When competitive interactions change, the relative sizes of the different stable regions change as well. Each of these different regions corresponds to the execution of a qualitatively different perception-action behavior.

What are the special requirements of systems that must interact with the physical world and also reason and solve problems? It is this question that must be addressed before we can claim a theory of intelligent physical agents and before AI and robotics can be reunited.

Comparing Approaches

We tested the two systems described previously to evaluate their relative performance in autonomous navigation tasks (Large, Christensen, and Bajscy 1999). We vary environmental complexity by constructing environments with different numbers of obstacles arranged in various configurations. We also vary task complexity; for example, a single agent performs simple navigation, or a pair of agents performs a cooperative task. In a range of tests, dynamic agents perform tasks faster and more reliably than discrete-event agents. They are able to maintain higher mean velocities, finding targets faster and with lower failure rates than discrete-event agents. However, the discrete-event agents also have advantages. They obey task constraints more faithfully, reluctant to relax constraints regardless of environmental complexity. These differences are the result of the model of decision making. The symbol-level dynamic system changes state less often than the discrete-event system, especially in the face of a noisy sensor reading. Although our experiments are not yet conclusive, by comparing modeling strategies in a careful way, we are gaining important insights into the special requirements of systems that must manipulate both signals and symbols at the same time and toward the same goal.

Conclusions

Early in the history of AI, many researchers came to believe that perception and action could be modeled by relatively simple transduction mechanisms, and therefore, abstract reasoning and problem solving were the difficult issues worthy of study. More recently, it has been argued that complex representation and reasoning might be unnecessary because many apparently intelligent behaviors can be modeled as perception-action systems situated in the physical world. Unfortunately, we have come to view both points of view as somewhat simplistic. Intelligent agents must be capable of bringing to bear a rich variety of perception-action strategies but, at the same time, reasoning and solving problems to perform both familiar and unfamiliar tasks in novel environments.

In this regard, the study of intelligent physical agents and their behavior is of tremendous theoretical and practical significance in AI. Not only are there a vast number of real-world applications where autonomous agents can be useful, but models of intelligent physical agents can serve as valuable starting points for theories of intelligent biological systems. The question that we ask is how to integrate models of perception-action behavior with models of problem-solving behavior.

Although we do not yet have the answer, we have two requirements for any solution: First, the description of the physical agent should take place within a structured framework that supports both analysis and theory making. Thus, we are allowed to develop design methodologies for artificial agents as well as develop rational theories of biological agents. Furthermore, any methodology that we propose should be compositional, allowing manageable and flexible system design through decomposition of complex problems or behaviors into subparts. Both of our systems meet these requirements. Finally, it is necessary to carefully compare the assumptions brought to bear by different strategies as we learn to model intelligent behavior in the real world.

What are the special requirements of systems that must interact with the physical world and also reason and solve problems? It is this question that must be addressed before we can claim a theory of intelligent physical agents and before AI and robotics can be reunited.

References

Alami, R.; Chatila, R.; Fleury, S.; Ghallab, M.; and Ingrand, F. 1998. An Architecture for Autonomy. *International Journal of Robotics Research* 17(4): 315–337.

Arkin, R. C. 1998. Behavior-Based Robotics. Cambridge, Mass.: MIT Press.

Brockett, R. W. 1993. Hybrid Models for Motion Control Systems. In *Essays in Control: Perspectives in the Theory and Its Applications*, eds. H. L. Trentelman and J. C. Willems, 29–53. Boston: Birkhauser.

Clementini, E.; Di Felice, P.; and Hernandez, D. 1997. Qualitative Representation of Positional Infor-

mation. *Artificial Intelligence* 95(2): 317–356.

Cohn, A. G. 1995. The Challenge of Qualitative Spatial Reasoning. *Computing Surveys* 27(3): 323–327.

Dickmanns, E. D. 1997. Vehicles Capable of Dynamic Vision: A New Breed of Technical Beings? In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Koenig, S., and Simmons, R. G. 1998. XAVIER: A Robot Navigation Architecture Based on Partially Observable Markov Decision Process Models. In *AI and Mobile Robots*, eds. D. Kortenkemp, R. P. Bonasso, and R. Murphy. Cambridge, Mass.: MIT Press.

Kosecka, J. 1996. A Framework for Modeling and Verifying Visually Guided Agents: Design, Analysis, and Experiments, GRASP TR 402, Ph.D. dissertation, School of Engineering and Applied Science, University of Pennsylvania.

Large, E. W.; Christensen, H. I.; and Bajscy, R. 1999. Scaling the Dynamic Approach to Path Planning and Control: Competition among Behavioral Constraints. *International Journal of Robotics Research* 18(1): 37–58.

Nagel, H. H., and Haag, M. 1998. Bias-Corrected Optical Flow Estimation for Road Vehicle Tracking. In *Proceedings of the International Conference on Computer Vision (ICCV '98)*, 1006–1011, New Delhi, India: Narosa.

Nerode, A., and Remmel, J. B. 1991. A Model for Hybrid Systems. Paper presented at the Hybrid Systems Workshop, 17–19 May, Ithaca, New York.

Ramadge, P. J., and Wonham, W. M. 1987. Supervisory Control of a Class of Discrete Event Processes. *SIAM Journal on Control and Optimization* 25(1): 206–230.

Russell, S. J., and Subramanian, D. 1995. Provably Bounded-Optimal Agents. *Journal of AI Research* 2(1): 575–609.

Sandewall, E. 1994. *Features and Fluents: The Representation of Knowledge about Dynamical Systems.* Oxford, U.K.: Oxford University Press.

Schoner, G.; Dose, M.; and Engels, C. 1996. Dynamics of Behaviour: Theory and Applications for Autonomous Robot Architectures. *Robotics and Autonomous Systems* 16(4): 213–246.

Shi, J., and Malik, J. 1998. Self-Inducing Relational Distance and Its Application to Image Segmentation. In *Proceedings of the 1998 European Conference on Computer Vision, Volume 1*, eds. H. Burkhardt and B. Neumann, 528–543. Lecture Notes in Computer Science. Berlin: Springer Verlag.

Steinhage, A., and Schoner, G. 1998. Dynamic Systems for the Behavioral Organization of Autonomous Robot Navigation. In *Sensor Fusion and Decentralized Control in Robotic Systems: Proceedings of the International Society for Optical Engineering*, eds. P. S. Schenker and G. T. McKee, 169–180. Bellingham, Wash.: International Society for Optical Engineering.

Tari, S.; Shah, J.; and Pien, H. 1997. Extraction of Shape Skeletons from Gray-Scale Images. *Computer Vision Image Understanding* (Special Issue on Biomedical Image Analysis) 66:133–146.

Ruzena Bajcsy obtained her first Ph.D. in EE from Slovak Technical University (Czechoslovakia) in 1967; she was the first woman in Slovakia ever to obtain a Ph.D. She then went to Stanford University and obtained her second Ph.D. in 1972, studying AI under John McCarthy. Her dissertation topic was machine perception, and she wrote one of the first programs enabling recognition of textured patterns.

She then joined the faculty at the University of Pennsylvania, where she has continued her work on machine perception and computer vision, characterizing and solving problems involving segmentation, three-dimensional vision, and other sensory modalities that function together with vision (for example, touch). Bajcsy's General Robotic Active Sensory Perception (GRASP) Laboratory at the University of Pennsylvania is internationally recognized in the field.

She chaired her department from 1984 through 1986 and has served on numerous National Research Council and National Science Foundation (NSF) advisory boards and committees. She was elected a fellow of the Institute of Electrical and Electronics Engineers in 1992 and the Association of Computing Machinery in 1995; she is also a founding member of the American Association for Artificial Intelligence. In 1995, she was elected as a member of the National Institute of Medicine, and she became a member of the National Academy of Engineering in 1997. She has served for three years on the CRA board.

Bajcsy is currently the assistant director of the NSF Directorate for Computer and Information Science and Engineering. Her e-mail address is bajcsy@central.cis.upenn.edu.

Edward Large is an assistant professor at the Center for Complex Systems at Florida Atlantic University. He received his Ph.D. from The Ohio State University in 1994. He has held fellowships in cognitive science and psychology at the University of Pennsylvania and in AI at Toshiba's Research and Development Laboratory in Kawasaka, Japan. His research focuses on the dynamics of human perception, action, and cognition and the design of autonomous agents using dynamic principles. His areas of specialization include music perception, auditory perception, and robotics. His e-mail address is large@walt.ccs.fau.edu.

Catadioptric Sensors that Approximate Wide-angle Perspective Projections

R. Andrew Hicks
Department of Mathematics and Computer Science
Drexel University
ahicks@drexel.edu

Ruzena Bajcsy
National Science Foundation
CISE Directorate
rbajcsy@nsf.gov

Abstract

We present two families of reflective surfaces that are capable of providing a wide field of view, and yet still approximate a perspective projection to a high degree. These surfaces are derived by considering a plane perpendicular to the axis of a surface of revolution and finding the equations governing the distortion of the image of the plane in this surface. We then view this relation as a differential equation and prescribe the distortion term to be linear. By choosing appropriate initial conditions for the differential equation and solving it numerically, we derive the surface shape and obtain a precise estimate as to what degree the resulting sensor can approximate a perspective projection. Thus these surfaces act as computational sensors, allowing for a wide-angle perspective view of a scene without processing the image in software. The applications of such a sensor should be numerous, including surveillance, robotics and traditional photography.

Recently, many researchers in the robotics and vision community have begun to consider visual sensors that are able to obtain wide fields of view. Such devices are the natural solution to various difficulties encountered with conventional imaging systems.

The two most common means of obtaining wide fields of view are fish-eye lenses and reflective surfaces, also known as catoptrics. When catoptrics are combined with conventional lens systems, known as dioptrics, the resulting sensors are known as **catadioptrics**. The possible uses of these systems include applications such as robot control and surveillance. In this paper we will consider only catadioptric based sensors. Often such systems consist of a camera pointing at a convex mirror, as in figure (1).

How to interpret and make use of the visual information obtained by such systems, e.g. how they should be used to control robots, is not at all obvious. There are infinitely many different shapes that a mirror can have, and at least two different camera models (perspective and orthographic projection) with which to combine each mirror.

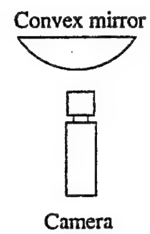

Figure 1. The generic setup of the type of sensor that we consider in this paper.

The properties of the resulting sensors are very sensitive to these choices.

The classic need for wide angle lenses has, of course been in photography. In particular, underwater and architectural photography are two examples in which having a wide-angle lens is often crucial. The commercially available lens with the widest field of view (without radial distortion) that the authors are aware of is the Nikon 13mm f/5.6 Nikkor AIS, which provides a field of view of 118 degrees at a cost of \$(US)12000. Note that our prototype orthographic sensor provides a field of view of 142 degrees.

1 Related work

In the past few years, there has been a tremendous increase in research on the design and applications of catadioptric based sensors. Much of this work has been focused on designing sensors with a panoramic or wide field of view (see [16], [10], [13], [17], [9], [7] [11],[18] [3], [19], [15], [12], [14], [4], [8],[1], [6]).

In [14], Nayar describes a true omni-directional sensor. In this case, the goal was to reconstruct perspective views. This sensor uses a parabolic mirror, which is essentially the only shape from which one can do a perspective unwarping of the image when using a camera that is well modeled by an orthographic projection (see [1]).

A different use of catadioptric sensors is an application

of Mouadddib and Pegard [15]. In this case a conical mirror is used to estimate a robot's pose. This is done using vertical lines in the world as landmarks, which appear as radial lines in the image. If the positions of these landmarks are known, then they can be used to estimate the robot's pose. In contrast to [14], in this work the authors use their device as a 2D sensor. The effect of noise on computing a robot's position by measuring the angles between known landmarks is investigated in [2]. Navigation and map building with a mobile robot using a conical mirror is considered by Yagi et al in [19] and [18].

In [4], Chahl and Srinivasan describe a means of estimating range by moving a panoramic sensor, based on the fact that the local distortion of the image is range dependent. This method, which gives a range estimate in every azimuthal direction, is implemented using a conical mirror.

The work most related to this paper is [6] and [5]. In [6], Conroy and Moore derive a family of mirrors for which the resolution in the image is invariant to changes in elevation. In [5] the authors exhibit a family of reflective surfaces that preserve a linear relationship between the angle of incidence of light onto a surface and the angle of reflection onto the imaging device.

Finally, we should point out that a preliminary form of this work is [8]. In this work the very same mirrors are discussed that we consider here. There is a major difference though, namely that at that point the authors were unaware that these mirrors approximated perspective projections in general, i.e. we believed that they only would unwarp the single plane that they were modeled on. It was after this work that the authors noticed this property experimentally and proved the approximation that appears below.

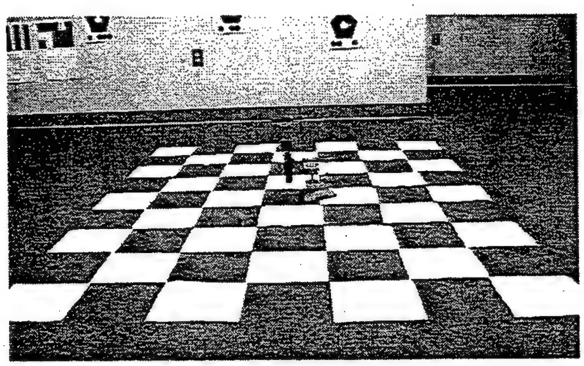

Figure 2. Here we see a panoramic sensor on a floor surrounded by 8.5 inch square sheets of paper.

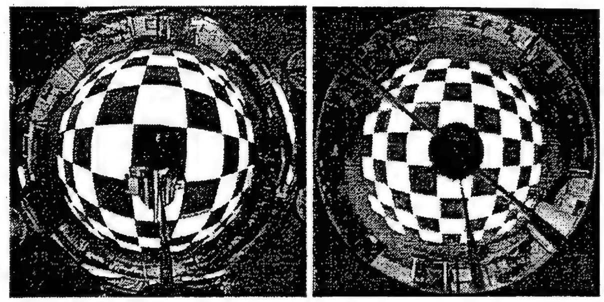

Figure 3. On the left we see a checkerboard scene similar to the one in figure (2), but now from the viewpoint of a catadioptric sensor consisting of a spherical mirror and a standard camera that gives an approximate perspective projection. On the right the viewpoint of a sensor that uses a parabolic mirror coupled with a camera that gives an approximate orthographic projection. (This system was purchased from Cyclovision Inc.) In each case, the mirrors are approximately thirty centimeters above the ground. Notice how the size of the squares decreases as a function of their distance from the camera and that the distortion caused by the spherical mirror is much greater than the distortion caused by the parabolic mirror.

2 Contributions

In figure (2) we see a scene consisting of a checkerboard pattern spread out on the floor around a panoramic sensor. Images of this scene (and different, but similar scenes) taken from sensors using spherical and parabolic mirrors appear in figure (3). It is clear that the distortion caused by the spherical mirror is greater than that caused by the parabolic mirror.

In this paper we present a class of sensors that provide a wide field of view with a perspective-like projection without any processing in software. In particular, it is possible to create a mirror that does not distort the checkerboard at all, other than by a chosen scaling factor. Images taken from such a sensor appear in figures (4) and (5). The key to finding the shape of this mirror is to write down the relationship between the equation of the mirror and how it distorts the checkerboard. This equation will contain the derivative of the function describing a cross section of the mirror, and may be considered as a means for finding the distortion if the mirror shape is given. On the other hand it can be considered as a differential equation in the shape of the mirror if the distortion function is given. By prescribing the

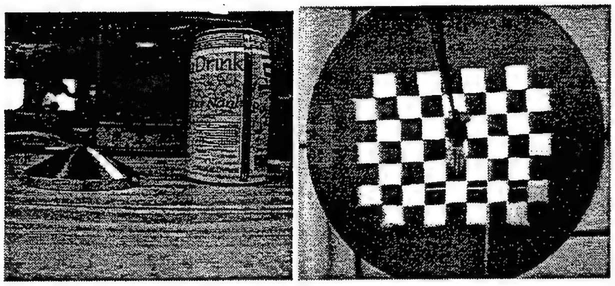

Figure 4. On the right we see a checkerboard scene (like that in figure (2), but with fewer checkers) from the viewpoint of a sensor that employs a new type of mirror. The shape of this mirror was determined by numerically solving a non-linear differential equation. In this image, the mirror is approximately thirty centimeters above the checkerboard pattern, just as the spherical and parabolic mirrors were in figure (3). On the left we see the mirror that was used to obtain this image.

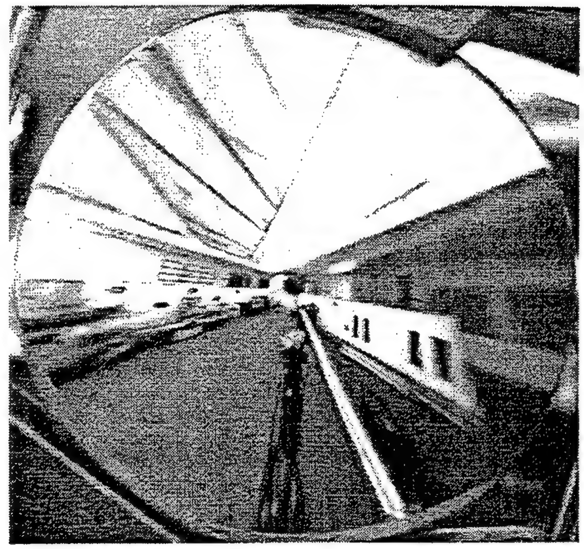

Figure 5. An image taken with a catadioptric sensor employing a pinhole mirror.

distortion to be linear and solving the differential equation numerically, data points describing the cross section were generated, which were then used to make prototypes out of steel or aluminum on a CNC lathe or mill. The cost of making each of these prototypes was approximately $ (US) 700.00.

3 Prescribing the Distortion

In this section we derive the an equation that leads to the construction of one of the two different types of mirrors. One model is based on the perspective projection and the other based on an orthographic projection. For this reason we will refer to the two mirrors as the "pinhole mirror" and the "orthographic mirror". The pinhole mirror is more natural in the sense that the pinhole camera is a good model for the imaging devices used in many applications. On the other hand, an orthographic projection is not difficult to achieve using the appropriate optics, and the mathematics associated with it is often simpler than for the pinhole model. We will omit the perspective model for reasons of space.

It is clear from figure (3) that if an object is on the floor, then the planar distance from the optical axis of the camera to any visible point of the object that touches the floor is a monotonic function of the pixel distance in the image. Therefore we have a distance function *d*, which takes pixel distances in the image and returns real world distances in the plane. It can also be seen from figure (3) that the distance function is rapidly increasing and approaches infinity as the horizon line is approached. It seems natural then to consider whether the mirror shape could be altered so that the distance function would be linear (or some other desired function), as is indicated in figure (6). We will refer to any such mirror with a linear distance function as a **rectifying mirror**.

We begin by deriving the equation for *d* for an arbitrary mirror. Consider an orthographic camera pointing up at a curved mirror, as is schematically depicted in figure (7). Here we see a cross section of the system, which is all we need consider since the mirror is rotationally symmetric. Our goal is to find an expression for $d(x)$ given the equation of the cross section of the mirror, F, and a point whose distance from the optical axis is x in the image plane.

From the diagram we have that $\tan(\theta) = F'(x)$, so that $\tan(2\theta) = \frac{2F'(x)}{1-F'(x)^2}$. On the other hand the diagram implies that $\tan(2\theta) = \frac{d(x)-x}{F(x)}$. Thus we have the equation

$$\frac{2F'(x)}{1-F'(x)^2} = \frac{d(x)-x}{F(x)}. \tag{1}$$

There are two ways to view equation (1). The first is what we just described above, i.e. if one knows F, it may be substituted into the above equation to determine $d(x)$, which is how the curves in figure (6) were computed. On the other hand, one could *choose* $d(x)$ and then consider (1) to be a differential equation satisfied by F. If we solve this differential equation, then the resulting mirror will have *d* as its distance function. It is important to note though, that at this point we know only that this property holds only in the one chosen plane, and does not other parallel planes. The

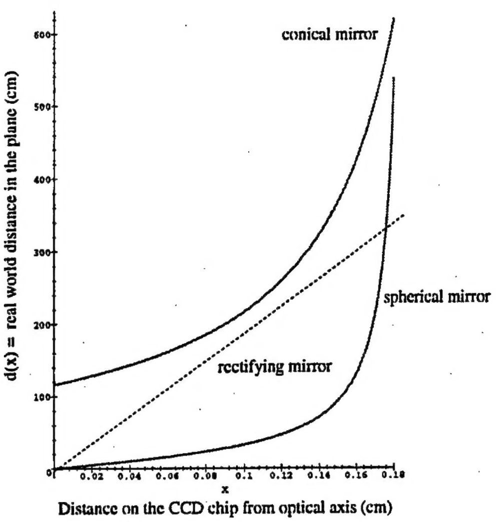

Figure 6. Plots of the distance functions for a conical mirror, a spherical mirror, and a rectifying mirrors, which we define as any mirror whose distance function is linear. Notice that the distance function for the conical mirror grows very linearly at first, but does not start at 0 cm, i.e. it omits a region of the plane surrounding the sensor.

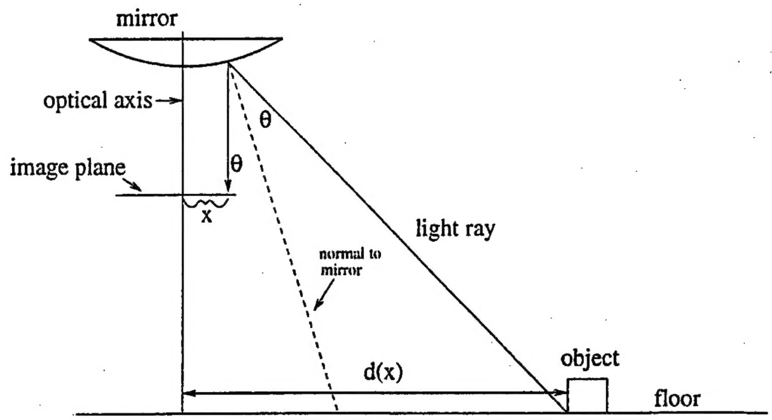

Figure 7. A schematic diagram of a catadioptric sensor that uses a camera modeled with an orthographic projection

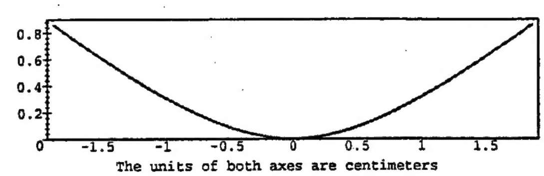

Figure 8. The cross section of the orthographic mirror used to create the image in figure (9).

reason for this is that the collection of rays that pass through two planes and reflect off of the mirror onto the image plane cannot be extended through the mirror to meet at a single point, i.e. the correspondence between the planes is not a perspective mapping. This point is discussed in more detail below.

A natural choice for the distance function is $d(x) = ax$ in equation (1). Considered as an differential equation, (1) is non-linear and numerical methods are called for. For our prototype orthographic mirror we chose a radius of 1.85 centimeters, $d(x) = 54x$ and the initial value $F(0) = 34$. Hence the field of view was $\arctan((54 \cdot 1.85)/34)) \sim 142$ degrees. The resulting cross section can be seen in figure (8).

In figure (9) we see an image taken using a mirror created from the cross-section depicted in figure (8). (As is evident from this example, the prototypes were of low optical quality, and the authors hope to have superior versions of them made in the future.) The images presented earlier, in figures (4) and 5) are from a perspective mirror.

4 Approximating Perspective Projections

Our above model was derived by considering how the sensor transformed a single plane, which we will always refer to as the **floor**. While for both the pinhole mirror and the orthographic mirror it is clear from experiments and simulations that planes perpendicular to the optical axis will be scaled by a constant, it is possible to show mathematically that with the properly chosen parameters, these mirrors will actually approximating a perspective projection to a high degree. In this section we derive the approximation for the orthographic mirror.

At first it may appear that a rectifying mirrors should only scale planes and not distort them, but in fact they do both distort them a little. In order to see how an arbitrary plane, P, is imaged, we need to know how P is mapped to the floor. If it is to be proportionally imaged, i.e. only transformed by a scale factor, then it must be transformed onto the floor by a scaling factor. For this to occur, the light rays that are entering the sensor by reflecting off of the mirror must have the property that, when extended beyond the mirror, they all intersect in a common point (see figure

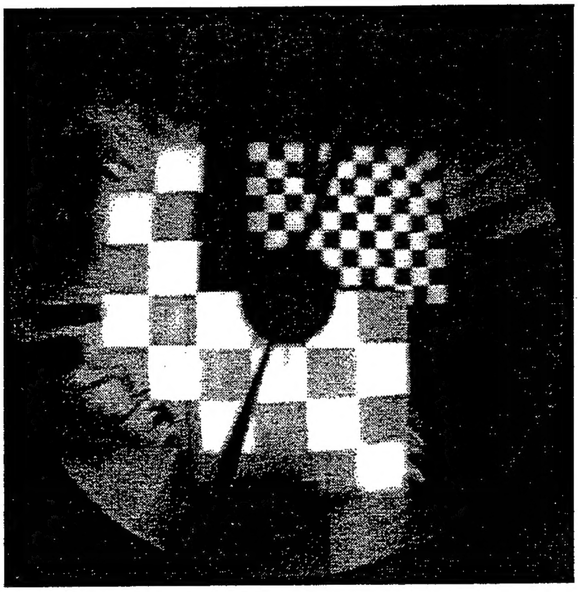

Figure 9. The image of two collections of checkers from two different distances. Here the sensor was placed on the edge of a table top, and one pattern of checkers was placed around its base on the table while another lay on the floor next to the table.

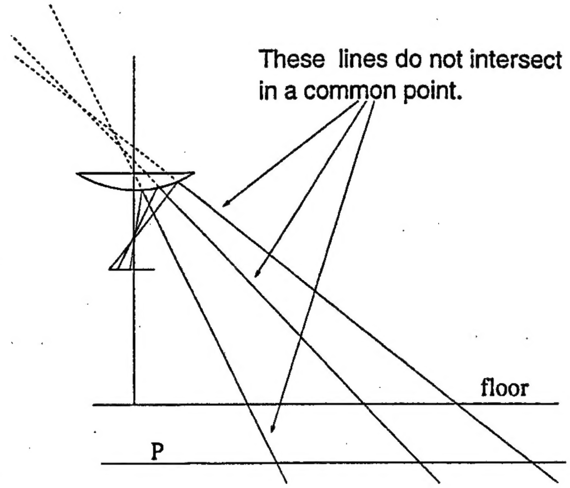

Figure 10.

(10)), i.e. there needs to be a single "effective viewpoint" for the sensor. Such a point does not exist for our two types of sensors, because, as is shown in [14], the only two catadioptric sensors with this property are the parabola coupled with an orthographic projection and the hyperbola coupled with a pinhole projection. Finally, we know that our surfaces are not paraboloids and hyperboloids, because as can be checked, parabolas and hyperbolas are not solutions to the appropriate differential equations.

We can demonstrate why the orthographic mirror gives a perspective-like projection by computing exactly how points in P are scaled onto the floor. In figure (11) we see that a point in P with distance d' from the optical axis is mapped to a point of distance $d(x)$, which in turn is mapped to a point in the image plane of distance x from the optical axis. We suppose that the distance between the two planes is r and the height of the mirror is $F(0)$. Then clearly

$$\frac{F(x)}{d(x)-x} = \frac{r}{d'-d(x)}, \tag{2}$$

from which it follows that

$$d'(x) = \frac{rd(x) - rx + F(x)d(x)}{F(x)} \tag{3}$$

Recall that $d(x) = \alpha x$ in our model, where generally we think of alpha as large (α is 54 for our prototype). Hence

$$d'(x) = \frac{(r - \frac{r}{\alpha} + F(x))d(x)}{F(x)} \tag{4}$$

Since we generally consider alpha to be large, we have the approximation

$$d'(x) = \frac{r + F(x)}{F(x)} d(x) \tag{5}$$

For our prototype orthographic mirror, $F(0)$ was chosen to be 34 cm, and x varied from 0 to 1.85 cm. The maximum value of F occurs at $x = 1.85$ cm with $F(1.85) = 34.85$ cm. Thus for that mirror $F(0) \sim F(x)$, which gives our final approximation:

$$d' \sim \frac{r + F(0)}{F(0)} d(x). \tag{6}$$

This last equation implies that this particular orthographic mirror will approximate a perspective projection with a pinhole placed at $(0, F(0))$.

Finally, for the purpose of comparison, it would be ideal if it was possible to compare images obtained from the orthographic mirror with an image taken with a lens providing the same field of view. As mentioned above, the widest possible field that can be obtained from a commercially available lens is 118 degrees. Thus, a simulation is a reasonable way to perform such a comparison. For this we used the Persistence of Vision raytracer, which will simulate a pinhole camera with any given field of view[1].

Consider the image in figure (12). In this scene we see a desk with some familiar objects. A pair of spheres are floating above the desk and to the right of the table a portion of a large box is visible. There are a number of other

[1]The desk scene by Tom Price and Dan Farmer and the teapot by Alexander Enzmann are both POV-Ray sample files.

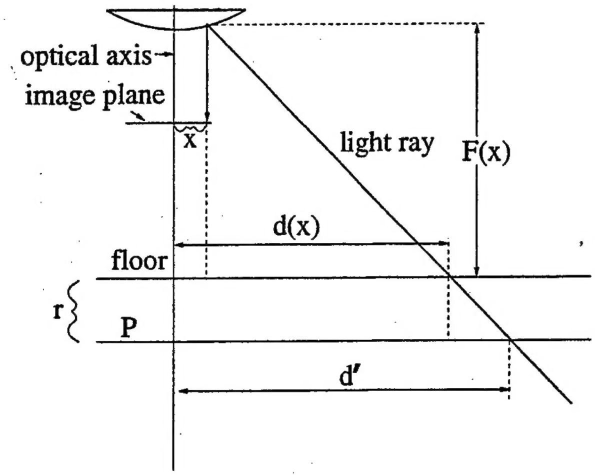

Figure 11. We consider a single ray of light passing through a plane P and then the floor. The goal is to derive an approximation to show that planes other than the floor map onto the floor by scaling.

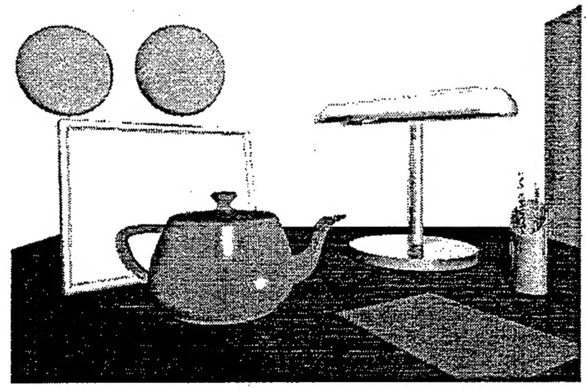

Figure 12. Here we see a scene created with the POV-Ray raytracer. In addition to the familiar objects on the desks are many blue spheres, only two of which are visible in this image. To the right of the desk can be seen the edge of a large box. The camera has a field of view of 67 degrees.

blue spheres to the left of the two that are visible, but these cannot be seen because the field of view is only 67 degrees. This image is formed by using a perspective projection.

In figure (13) we see the same scene imaged with the orthographic mirror. The mirror has been placed in the same position as the camera was when figure (12) was created. The table top objects on the table are relatively undistorted, as is the box, now fully visible on the left. Since the field of view has been increased, we can now see all of the blue spheres and the entire box.

In figure (14) we see the same desk scene again, this time with a simulated camera whose field of view is 142 degrees. These two images agree to a great extent, with the obvious exception of the reflection through the vertical axis, and demonstrate the fineness of the approximation.

5 Conclusion

We have exhibited a sensor design which has the ability to give a normal camera an ultra-wide field. These sensors are based on a family of mirrors derived as numerical solutions of non-linear differential equations which describe how a plane perpendicular to the optical axis of the system is distorted. By using the geometry of the mirror, the image is unwarped in an analog manner, and so requires no processing time, and thus these devices may be considered as "computational analog sensors". These sensors could be useful for applications such as human monitored surveillance systems and would not require a digital computer. In addition, if one did have a computer available, applications such a motion detection are simplified since the sensor will provide a uniform resolution image of the floor of a room, as opposed to a spherical or parabolic mirror in which the outer regions would suffer from low resolution.

Acknowledgments

The authors would like to thank Stephen Shatz and David Pettey for many useful discussions about this subject.

References

[1] S. Baker and S. Nayar. A theory of catadioptric image formation. In *Proc. International Conference on Computer Vision*, pages 35–42, 1998.

[2] M. Betke and L. Gurvits. Mobile robot localization using landmarks. *IEEE Transactions on Robotics and Automation*, 13:251–263., 1997.

[3] S. Bogner. Introduction to panoramic imaging. In *Proceedings of the IEEE SMC Conference*, pages 3100–3106, 1995.

[4] J. Chahl and M. Srinivasan. Range estimation with a panoramic sensor. *J. Optical Soc. Amer. A*, 14:2144–2152, 1997.

[5] J. Chahl and M. Srinivasan. Reflective surfaces for panoramic imaging. *Applied Optics*, 36:8275–8285, 1997.

[6] T. Conroy and J. Moore. Resolution invariant surfaces for panoramic vision systems. In *Proc. International Conference on Computer Vision*, pages 392–397, 1999.

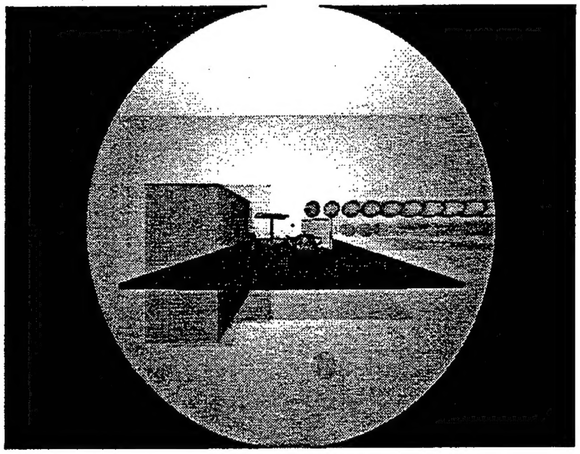

Figure 13. Here we see the same scene as in figure (12), imaged with the orthographic mirror. The mirror was place in the same position as the camera in figure (12), but much more of the scene is now visible. Flat objects, such as the large box are relatively undistorted, as are the objects on the table. The spheres though become increasingly distorted as they approach the edge of the image. This is the so-called elliptical distortion of perspective projection.

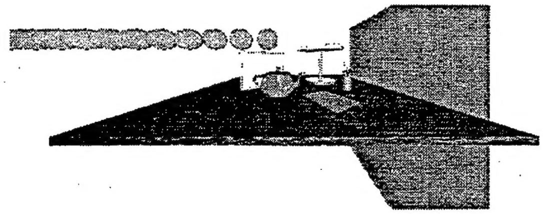

Figure 14. Here we see the same scene as in figure 12, imaged with pinhole camera whose field of view is 142 degrees (the same as the orthographic mirror). The camera is in the same position as in the previous two figures. Thus this simulation is consistent with the mathematics showing that the orthographic mirror gives a good approximation to a perspective projection.

[7] A. Goshtasby and W. Gruver. Design of a single lens stereo camera system. *Pattern Recognition*, 26:923–937, 1993.

[8] R. Hicks. Reflective surfaces as computational sensors. In *Proceedings of the 2nd Workshop on Perception for Mobile Agents, CVPR 99*, pages 82–86, 1999.

[9] J. Hong. Image based homing. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 1991.

[10] C. S. J. Charles and R. Reeves. *How to build and use an all sky camera.* Astronomy Magazine, April, 1987.

[11] Y. Y. K. Yamazawa and M. Yachida. Omnidirectional imaging with hyperboidal projection. In *Proceedings of the IEEE International Conference on Robots and Systems*, 1993.

[12] V. Nalwa. A true omnidirectional viewer. In *Technical Report, Bell Laboratories, Holmdel, NJ 07733, USA*, 1996.

[13] S. Nayar. Sphereo: Recovering depth using a single camera and two specular spheres. In *Proceedings of SPIE: Optics, Illumination, and Image Sensing for Machine Vision II*, pages 251–263, 1988.

[14] S. Nayar. Catadioptric omnidirectional camera. In *Proc. Computer Vision Pattern Recognition*, pages 482–488, 1997.

[15] C. Pegard and E. Mouaddib. A mobile robot using a panoramic view. In *Proc. IEEE Conference on Robotics and Automation*, pages 89–94, 1996.

[16] D. Rees. *Panoramic television viewing system.* United States Patent No. 3,505,465, April, 1970.

[17] Y. Yagi and S. Kawato. Panoramic scene analysis with conic projection. In *Proceedings of the International Conference on Robots and Systems*, 1990.

[18] Y. Yagi, S. Kawato, and S. Tsuji. Real-time omnidirectional image sensor (copis) for vision-guided navigation. *Trans. on Robotics and Automation*, 10:11–22, 1994.

[19] Y. Yagi, S. Nishizawa, and S. Tsuji. Map-based navigation for a mobile robot with omnidirectional image senso. *Trans. on Robotics and Automation*, 11:634–648, 1995.